Noise Suppression and Spectral Decomposition for State-Dependent Noise in the Presence of a Stationary Fluctuating Input

D. Brian Walton[1] and Koen Visscher[2]

[1]*Department of Applied Mathematics, University of Washington, Seattle, Washington 98195*

[2]*Departments of Physics and Molecular and Cellular Biology, University of Arizona, Tucson, Arizona 85721*

(Dated: 29 August 2003)

Abstract

It recently has been shown that the observed noise amplitude of an intrinsically noisy system may be reduced by causing the underlying state to fluctuate [J. M. G. Vilar and J. M. Rubí, Phys. Rev. Lett. **86**, 950 (2001)]. This paper extends the previous theory by considering the full power spectrum of the output signal, interpreting noise reduction in terms of the low-frequency end of the spectrum as well as the integrated spectrum. While the former provides a measure of the variance in the estimator of the mean at asymptotically long time scales, the latter provides the variance of discretely sampled observations. Our treatment accounts for arbitrarily sized fluctuations and deals with both continuous and discretely sampled observations. We show that noise suppression is possible if and only if the stationary average of the intensity of state-dependent noise decreases. We apply our analysis to a an example involving saturable electrical conduction discussed in the original paper by Vilar and Rubí.

PACS numbers: 05.40.Ca, 05.45.Tp, 02.70.Hm

Keywords: Noise, Ergodic Average, Power Spectrum, Hermite Expansion

I. INTRODUCTION

Although noise is generally considered a nuisance, its constructive properties have become increasingly clear in recent years as it plays essential roles in many fundamental processes, particularly in biology. For example, biological motion by motor proteins requires the presence of thermal fluctuations [1, 2]. Furthermore, the response of sensory systems can be enhanced through the interactions of noise and a weak driving signal leading to stochastic resonance [3–5]. Recently, another beneficial aspect of noise has come to light: noise may be used to reduce noise [6].

In their letter "Noise Suppression by Noise" [6], Vilar and Rubí consider the effect of adding fluctuations to the input of a system where the intrinsically noisy output signal depends on the input signal. In particular, they propose a model where, for a steady input, the rapid fluctuations in the output signal due to intrinsic noise are essentially white, with an intensity that depends on the particular input signal. Assuming that the autocorrelation time of the noise added at the input is short compared to times of interest, the output signal has a spectrum that remains essentially white over the low frequencies of interest. Vilar and Rubí demonstrate that the spectral intensity of the output noise at these low frequencies can be reduced by adding fluctuations at the input of the system. They also give a sufficient condition for such a reduction when the added fluctuations at the input are kept sufficiently small.

In this paper, we provide a more complete mathematical basis for the phenomenon of noise suppression. In particular, we provide a framework for predicting the power spectral density for arbitrary scales of input fluctuations. We also adapt the original model to account for discretely sampled observations. The overall fluctuations in the output signal naturally decouple into contributions arising from the intrinsic noise as well as the changing input. We take advantage of this decomposition to compute the spectral power from these two contributions separately. A consequence of this result is that noise suppression, which corresponds to a decrease in the spectral power of the signal at either specific frequencies or over a range of frequencies, will only occur if the average intensity of intrinsic noise decreases.

The article is organized as follows. First, we provide a formal description of the model describing the input and output signals and introduce the appropriate quantitative measures of noise. Next, we consider the power spectrum for the output signal, showing that it decomposes into a white component corresponding to the average intensity of intrinsic noise plus the spectrum characterizing the output signal in the absence of intrinsic noise. We further demonstrate an explicit approach to determining this profile using an eigenmode expansion involving Hermite polynomials. Third, we discuss the possibility of noise suppression in the context of this spectral decomposition. Subsequently, we apply the techniques to one of the examples originally provided by Vilar and Rubí. We conclude by commenting on the basic mechanism required for noise suppression.

II. MODEL DESCRIPTION

The system under consideration essentially models an observed signal that is regulated by an underlying state, X, which could represent either an unobserved, internal dynamic variable or an external, controlled input signal. The output signal, Y, will be intrinsically noisy in that there will be uncorrelated fluctuations for every state of the input. We seek a model that captures the following properties. First, the output signal relaxes rapidly to an essentially stationary process that depends on the current value of the input signal, characterized by a mean signal level and by the intensity of the intrinsic, uncorrelated noise. We idealize this relaxation by assuming that Y instantaneously reflects the current state X. Second, the input signal will correspond to a stationary Gaussian process whose autocorrelation decays exponentially in time.

Let X_t and Y_t represent the value of the signals at the time t. The input signal X_t is modeled as a stationary, continuous Gaussian process with mean x_0, variance σ^2, and autocorrelation time τ. For a finite, arbitrary collection of times, $t_1 < t_2 < \cdots < t_n$, the distribution of X_t is characterized by the density

$$\Phi(t_1, x_1; t_2, x_2; \ldots; t_n, x_n) = \phi(x_1; x_0, \sigma^2) \prod_{l=2}^{n} \phi(x_l; \mu_l, \sigma_l^2), \tag{1}$$

where the individual factors are expressed in terms of the densities of Gaussian random variables with means

$$\mu_l = x_0 + (x_{l-1} - x_0)e^{-(t_l - t_{l-1})/\tau} \tag{2}$$

and variances

$$\sigma_l^2 = \sigma^2(1 - e^{-2(t_l - t_{l-1})/\tau}) \tag{3}$$

through the parametrized density for a Gaussian random variable with mean μ and variance η^2

$$\phi(x;\mu,\eta^2) = \frac{1}{\sqrt{2\pi\eta^2}} \exp\left(-\frac{(x-\mu)^2}{2\eta^2}\right). \tag{4}$$

The autocovariance of X_t is given by $\gamma_X(u)$ as

$$\gamma_X(u) = \langle \widetilde{X}_t \, \widetilde{X}_{t+u} \rangle = \sigma^2 e^{-|u|/\tau}, \tag{5}$$

where $\widetilde{X}_t = X_t - x_0$ represents the centered state, and averages are with respect to the given distribution. This process is equivalent to the Ornstein-Uhlenbeck velocity process [7], but with a shifted mean. For discretely sampled observations of the output, we define observation times $\{t_k = k\Delta t;\ k \geq 0\}$ in terms of the sampling interval Δt, and for notational simplicity write $X_k = X_{t_k}$ and $Y_k = Y_{t_k}$. In this case, we naturally define the autocorrelation $\rho = e^{-\Delta t/\tau}$, and the sequence $\{X_k\}$ becomes a simple autoregressive process (often written AR(1)) with autocorrelation ρ and variance σ^2 [8], which has a corresponding discrete autocovariance function $\gamma_{X,\Delta}(p)$ for lag p as

$$\gamma_{X,\Delta}(p) = \langle \widetilde{X}_k \, \widetilde{X}_{k+p} \rangle = \sigma^2 \rho^{|p|}. \tag{6}$$

Setting $\sigma^2 = 0$ corresponds to a constant input signal.

The observation process Y_t will be a function of the state X_t with additive state-dependent white noise so that we write

$$Y_t = H(X_t) + g(X_t)\xi_t. \tag{7}$$

The process ξ_t represents a standard continuous-time white noise process that is independent of the input process X, having zero mean, $\langle \xi_t \rangle = 0$, and autocovariance

$$\langle \xi(t)\xi(s) \rangle = \delta(t-s). \tag{8}$$

The function $H(X)$ represents the mean output signal for a given state X, while the function $g(X)$ establishes the intensity of output noise for that state. If the state X remains constant ($\sigma^2 = 0$), then Y has an autocovariance given by

$$\langle (Y_t - H(X))(Y_s - H(X)) \rangle = G(X)\delta(t-s), \tag{9}$$

expressed in terms of the covariance intensity $G(X) = g^2(X)$. We remark that mathematically, the continuous white noise process ξ_t and consequently Y_t are not well-defined stochastic processes, but in fact should be expressed as stochastic differentials [7]. The white noise ξ_t corresponds to the differential of a Wiener process and Y_t corresponds to the differential of a diffusion with state-dependent infinitesimal drift $H(X)$ and variance $G(X)$. Nevertheless, the formal expressions given above will suffice for this paper. In the discretely sampled case, we introduce the discrete white-noise process $\{\varepsilon_k\}$, an independent and identically distributed sequence of Gaussian random variables each with zero mean and unit variance, which is independent of the state process $\{X_k\}$, so that we may write

$$Y_k = H(X_k) + \beta(X_k)\varepsilon_k. \tag{10}$$

We remark that the variance $\beta^2(X)$ arising the sampled case may or may not be related to an underlying continuous intensity $G(X)$. If we integrate the observation Y_t over an interval of duration Δt which is short compared to the autocorrelation time τ, then X_t remains essentially unchanged over that interval and we may approximate

$$\int_t^{t+\Delta t} Y_s\,ds \approx H(X_t)\Delta t + G(X_t)\Delta W_{\Delta t} \tag{11}$$

where $\Delta W_{\Delta t}$ is a Gaussian random variable with mean zero and variance Δt. Dividing both sides of this by the interval duration Δt, we average the signal. That is, the discretely sampled sequence generated by averaging the continuous signal,

$$Y_k = \frac{1}{\Delta t}\int_{t_{k-1}}^{t_k} Y_t\,dt, \tag{12}$$

would be well-approximated by

$$Y_k \approx H(X_k) + G(X_k)\frac{\Delta W_{\Delta t,k}}{\Delta t}. \tag{13}$$

In order for the variance of the noise term of this approximation to match the discretely sampled model given in Eq. (10), we must have

$$\beta^2(X) = \frac{1}{\Delta t}G(X). \tag{14}$$

The essential property for associating the discrete noise variance with the continuous noise intensity is that averaging the signal accumulates error over the entire interval Δt at an approximately constant rate $G(X_t)$. If, however, the discrete observation results from a single observation occurring at the end of the sampling interval, the intrinsic noise results either from a brief accumulation of error related to the time to make the observation or else from other sources of error. In both cases, the discrete scale of noise $\beta^2(X)$ becomes independent of the time between samples, Δt, so that a relationship between $\beta^2(X)$ and some $G(X)$ through Eq. (14) would not hold.

The asymptotic suppression of noise discussed in [6] refers to a decrease in the spectral intensity at low frequencies. To be explicit, if $\gamma_Y(u)$ represents the autocovariance function for the observation signal Y_t with lag u, then the power spectral density $h_Y(\omega)$ at the frequency ω can be expressed in terms of the Fourier transform of $\gamma_Y(u)$

$$h_Y(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \gamma_Y(u)\, e^{-i\omega u}\, du \tag{15}$$

as guaranteed by the Wiener-Khintchine theorem [8]. When $\sigma^2 = 0$, the input state does not fluctuate, and the spectral density has a constant level of $G(x_0)/2\pi$. When $\sigma^2 > 0$, the shape of the spectrum changes. As we show later, for reasonable functions, H, the spectral density will be smooth, so that at sufficiently low frequencies, corresponding to asymptotically long time scales, the spectral density will be approximately constant at the level

$$h_Y(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \gamma_Y(u)\, du. \tag{16}$$

Defining the mean signal $H_0 = \langle Y_t \rangle$ and the asymptotic intensity of noise $G_0 = 2\pi h_Y(0)$, the process Y_t is approximately equivalent (at large time scales and in distribution) to a second process $\widehat{Y}_t$ that is independent of the fluctuating input state and given by

$$\widehat{Y}(t) = H_0 + \sqrt{G_0}\, \widehat{\xi}_t, \tag{17}$$

where $\widehat{\xi}_t$ is another continuous-time white noise process independent of ξ_t. In this specific sense, noise is suppressed if $G_0 < G(x_0)$ since the approximating process $\widehat{Y}_t$ has smaller intensity noise than Y_t would have had if $\sigma^2 = 0$. Vilar and Rubí [6] provide a perturbation expansion of the integrated scale of noise which provides a sufficient condition for noise suppression that $G''(x_0) < 0$ so long as σ^2 and τ are sufficiently small, although they also provide some numerical examples demonstrating that this persists for larger values of σ^2 but still with a small autocorrelation time $\tau \ll 1$.

III. SPECTRAL DECOMPOSITION

We now broaden the discussion and consider the complete power spectrum of the observation process. In order to characterize the spectrum of the output signal, we first need to describe precisely the autocovariance of Y. Because $\langle \xi_t \rangle = 0$ is independent of X_t, the mean signal H_0 is simply the stationary average of $H(X_t)$,

$$\begin{aligned} H_0 &= \langle Y_t \rangle = \langle H(X_t) \rangle \\ &= \int_{-\infty}^{\infty} H(x)\phi(x; x_0, \sigma^2)\, dx. \end{aligned} \tag{18}$$

The centered output signal $\widetilde{Y}_t = Y_t - H_0$ determines the autocovariance of the signal,

$$\gamma_Y(u) = \langle \widetilde{Y}_t \widetilde{Y}_{t+u} \rangle. \tag{19}$$

By writing $\widetilde{H}(x) = H(x) - H_0$, we can explicitly see the contributions arising from the white noise and from the fluctuations in the input,

$$\gamma_Y(u) = \langle (\widetilde{H}(X_t) + g(X_t)\xi_t)(\widetilde{H}(X_{t+u}) + g(X_{t+u})\xi_{t+u}) \rangle \tag{20}$$

$$= \langle \widetilde{H}(X_t)\widetilde{H}(X_{t+u}) \rangle + \langle G(X_t) \rangle \delta(u), \tag{21}$$

where we again use the assumption that ξ_t is independent of the input X. Thus, the autocovariance decomposes into two terms. The first term characterizes the correlated variability in the output signal arising from the autocorrelated input signal. The second term characterizes the average uncorrelated fluctuations due to the intrinsic white noise. By linearity of the Fourier transform, the power spectrum must also decompose into two contributions,

$$h_Y(\omega) = h_H(\omega) + \frac{1}{2\pi} \langle G(X_t) \rangle, \tag{22}$$

the sum of the spectral density of the stationary process $\widetilde{H}(X_t)$, given by $h_H(\omega)$, and the constant spectral density given by the average intensity of the white noise, $\langle G(X_t) \rangle$.

We now turn to an analytic approach to determine the power spectral density $h_H(\omega)$. The Ornstein-Uhlenbeck process that governs the input signal is a Markov process, so we may consider the transition semigroup T_t on functions of X defined by

$$T_t[f](x) = \langle f(X_t) \rangle_x \tag{23}$$

$$= \int_{-\infty}^{\infty} f(y) k(y; x, t) dy, \tag{24}$$

where the average in the first equality is conditioned on the process starting at $X_0 = x$ and where $k(y; x, t)$ is the transition probability density for $X_t = y$ given $X_0 = x$. The transition kernel $k(y;x,t)$ is the Gaussian density with mean $x_0+(x-x_0)e^{-t/\tau}$ and variance $\sigma^2(1-e^{-2t/\tau})$. The autocovariance $\gamma_H(u)$ can be expressed in terms of the semigroup as

$$\gamma_H(u) = \langle \widetilde{H}(X_t) T_u[\widetilde{H}](X_t)\rangle. \tag{25}$$

The semigroup operator can also be expressed in terms of its infinitesimal generator $\mathcal{L}$ as $T_t = e^{t\mathcal{L}}$, where $\mathcal{L}$ is defined as the differential operator

$$\mathcal{L}[f](x) = \frac{1}{\tau}[-(x-x_0)f'(x) + \sigma^2 f''(x)], \tag{26}$$

which governs the Kolmogorov backward, or adjoint, equation for the Ornstein-Uhlenbeck diffusion process [7]. If we define the inner product between functions f_1 and f_2 as the integral of the product with respect to the stationary probability measure,

$$\begin{aligned}\langle f_1, f_2\rangle &= \langle f_1(X_t) f_2(X_t)\rangle \\ &= \int_{-\infty}^{\infty} f_1(x) f_2(x)\phi(x;x_0,\sigma^2)dx,\end{aligned} \tag{27}$$

then $\mathcal{L}$ is a self-adjoint operator in the Hilbert space defined by this inner product.

We expand the function $\widetilde{H}$ in terms of the eigenfunctions of $\mathcal{L}$. Eigenfunctions $f_\lambda(x)$ with eigenvalue λ will be determined from the equation

$$\mathcal{L}[f_\lambda](x) = \lambda f_\lambda(x) \tag{28}$$

which can be rewritten as

$$-(x-x_0)f_\lambda'(x) + \sigma^2 f_\lambda''(x) = \tau\lambda f_\lambda(x). \tag{29}$$

If we shift and rescale space using the substitution $z(x) = (x-x_0)/\sqrt{2\sigma^2}$, the eigenvalue equation becomes

$$f_\lambda''(z) - 2z f_\lambda'(z) - 2\tau\lambda f_\lambda(z) = 0, \tag{30}$$

which is well-known to have as solutions the Hermite polynomials, $f_\lambda(x) = H_n(z(x))$, provided $\lambda = -n/\tau$ for integer values of $n \geq 0$ [9]. The functions H_n are orthogonal,

$$\langle H_n(\widetilde{X}_t/\sqrt{2\sigma^2}) H_m(\widetilde{X}_t/\sqrt{2\sigma^2})\rangle = \delta_{n,m}\, 2^n n!, \tag{31}$$

and form a complete basis for the Hilbert space [9]. So for a square-integrable function $\widetilde{H}$ (i.e. $\langle \widetilde{H}^2(X_t)\rangle < \infty$), we create the Hermite expansion

$$\widetilde{H}(x) = \sum_{n=1}^{\infty} \frac{c_n}{\sqrt{2^n n!}} H_n(\frac{x-x_0}{\sqrt{2\sigma^2}}), \tag{32}$$

starting the sum at $n = 1$ because $\langle \widetilde{H}(X_t) \rangle = 0$. The coefficients are determined in the standard way by computing the inner product

$$c_n = \frac{1}{\sqrt{2^n n!}} \langle \widetilde{H}(X_t) H_n(\widetilde{X}_t / \sqrt{2\sigma^2}) \rangle. \tag{33}$$

We emphasize that the coefficients actually depend on σ^2, as the Hermite expansion changes when σ^2 changes. In terms of the rescaled variable z, we have $x = x_0 + \sqrt{2\sigma^2}\, z$ so that

$$c_n = \frac{1}{\sqrt{2^n n!}} \int_{-\infty}^{\infty} \widetilde{H}(x_0 + \sqrt{2\sigma^2}\, z) H_n(z) \phi(z; 0, 1)\, dz, \tag{34}$$

so that from this perspective, changing σ fundamentally changes the function for which we find the expansion. The variance of $\widetilde{H}(X_t)$ can be simply expressed as the sum of the squares of the coefficients $\{c_n : n \geq 1\}$,

$$\langle \widetilde{H}^2(X_t) \rangle = \sum_{n=1}^{\infty} c_n^2. \tag{35}$$

We are now in a position to compute the autocovariance and hence the power spectrum. Using the orthogonality property and the diagonal action of the semigroup operator on the eigenfunctions,

$$T_u[H_n(\frac{x - x_0}{\sqrt{2\sigma^2}})] = e^{-nu/\tau} H_n(\frac{x - x_0}{\sqrt{2\sigma^2}}), \tag{36}$$

we find that the autocovariance of $\widetilde{H}(X_t)$ is

$$\gamma_H(u) = \sum_{n=1}^{\infty} c_n^2 e^{-nu/\tau}. \tag{37}$$

Taking the Fourier transform, we determine the power spectrum $h_H(\omega)$ as

$$h_H(\omega) = \frac{1}{2\pi} \sum_{n=1}^{\infty} c_n^2 \frac{2\tau n}{n^2 + (\tau\omega)^2}, \tag{38}$$

which can further be represented in terms of the spectral profile

$$\widehat{h}_H(\omega) = \frac{1}{2\pi} \sum_{n=1}^{\infty} c_n^2 \frac{n}{n^2 + \omega^2} \tag{39}$$

through the rescaling

$$h_H(\omega) = 2\tau \widehat{h}_H(\tau\omega). \tag{40}$$

The power spectrum for the discretely sampled observation process can be similarly computed. The autocovariance of the discretely sampled observation sequence $\gamma_{Y,\Delta}(p)$ decomposes into

$$\gamma_{Y,\Delta}(p) = \gamma_{H,\Delta}(p) + \langle \beta^2(X_k) \rangle \delta_{p,0}, \tag{41}$$

analogous to the decomposition for the continuous time process. So we again focus on determining the autocovariance and spectrum for the sequence $\widetilde{H}(X_k)$. Knowing the autocorrelation coefficient ρ and the sampling interval Δt (or setting $\Delta t = 1$ if unknown), we can determine the corresponding autocorrelation time $\tau = -\Delta t / \ln \rho$ so that the discrete autocovariance function for $\widetilde{H}(X_k)$ will be given by

$$\begin{aligned}
\gamma_{H,\Delta}(p) &= \gamma_H(p\Delta t) && (42)\\
&= \sum_{n=1}^{\infty} c_n^2 e^{-n|p|\Delta t/\tau} && (43)\\
&= \sum_{n=1}^{\infty} c_n^2 \rho^{n|p|}. && (44)
\end{aligned}$$

We compute the power spectral density $h_{Y,\Delta}(\omega)$ over the frequencies $\omega \in (-\pi/\Delta t, \pi/\Delta t)$ as the Fourier series

$$h_{Y,\Delta}(\omega) = \frac{\Delta t}{2\pi} \sum_{p=-\infty}^{\infty} \gamma_{Y,\Delta}(p) e^{-i\omega p \Delta t}, \tag{45}$$

which will decompose into the spectrum for $\widetilde{H}(X_k)$ plus a white-noise component,

$$h_{Y,\Delta}(\omega) = h_{H,\Delta}(\omega) + \frac{\Delta t}{2\pi} \langle \beta^2(X_k) \rangle. \tag{46}$$

Outside of the interval, $(-\pi/\Delta t, \pi/\Delta t)$, the spectral densities vanishes. The spectral density coming from $\widetilde{H}(X_k)$ can then be written

$$h_{H,\Delta}(\omega) = \frac{\Delta t}{2\pi} \sum_{n=1}^{\infty} c_n^2 \frac{1 - \rho^{2n}}{1 + \rho^{2n} - 2\rho^n \cos \omega \Delta t}. \tag{47}$$

For the case that the sampled data arise from the continuous model, ρ is defined in terms of the autocorrelation time τ and the sampling interval Δt as $\rho = e^{-\Delta t/\tau}$ so that, in the limit as $\Delta t \to 0$, the spectrum for the sampled observations, Eq. (47), recovers the spectrum for the continuous signal, Eq. (38).

IV. NOISE SUPPRESSION

With the spectrum in hand, we return to the question of noise suppression. First, we consider the asymptotic noise level G_0 originally considered by Vilar and Rubí. Recall that G_0 corresponds to the intensity of white noise that that gives the same spectral density as

Y_t for asymptotically small frequencies. This intensity can be explicitly written

$$G_0 = 2\pi h_Y(0) = 2\tau \sum_{n=1}^{\infty} \frac{c_n^2}{n} + \langle G(X_t) \rangle. \tag{48}$$

The decomposition clearly shows that a necessary condition for noise suppression in this sense, $G_0 < G(x_0)$ is that the average intensity of white noise decrease,

$$\langle G(X_t) \rangle < G(x_0), \tag{49}$$

since the effect of the fluctuating signal coming from $h_H(0)$ raises the spectral density even higher. In fact, if we add the constraint on the autocorrelation time

$$\tau < \frac{G(x_0) - \langle G(X_t) \rangle}{2 \sum_{n=1}^{\infty} c_n^2/n}, \tag{50}$$

we have a necessary and sufficient condition for noise suppression. We stress that a fixed value of σ is implicit in this statement, since both $\langle G(X_t) \rangle$ and the coefficients c_n depend on σ.

In a similar way, we might consider a random sequence $\{\widehat{Y}_k\}$ which is independent of the state X, given by the analog of Eq. (17),

$$\widehat{Y}_k = H_0 + \beta_0 \widehat{\varepsilon}_k, \tag{51}$$

where $\widehat{\varepsilon}_k$ is a white noise sequence. Such a sequence will have white power spectrum which will correspond to the spectral density of $\{Y_k\}$ at low frequencies, $h_{Y,\Delta}(0)$, if β_0^2 is defined as

$$\beta_0^2 = \sum_{n=1}^{\infty} c_n^2 \frac{1+\rho^n}{1-\rho^n} + \langle \beta^2(X_k) \rangle. \tag{52}$$

Noise suppression in this asymptotic sense will occur when the stationary average variance of the intrinsic noise decreases, $\langle \beta^2(X_k) \rangle < \beta^2(x_0)$, analogous to the sufficient condition for the continuous case. The sequence of factors $\{\frac{1+\rho^n}{1-\rho^n} : n = 1, 2, \ldots\}$ is a decreasing sequence converging to 1 so that we actually obtain bounds on β_0^2 in terms of the variance $\langle \widetilde{H}(X_k) \rangle$ and the average variance $\langle \beta^2(X_k) \rangle$ as,

$$\langle \widetilde{H}^2(X_k) \rangle + \langle \beta^2(X_k) \rangle < \beta_0^2 < \frac{1+\rho}{1-\rho} \langle \widetilde{H}^2(X_k) \rangle + \langle \beta^2(X_k) \rangle, \tag{53}$$

which actually avoids the explicit computation of the coefficients $\{c_n\}$. Consequently, we obtain the stronger necessary condition that the sum of the average variance of the white noise $\langle\beta^2(X_k)\rangle$ and the stationary variance of the signal $H(X_k)$ must be less than $\beta^2(x_0)$. We also obtain the sufficient condition that

$$\frac{1+\rho}{1-\rho}\langle\widetilde{H}^2(X_k)\rangle + \langle\beta^2(X_k)\rangle < \beta^2(x_0). \tag{54}$$

Because we actually have access to the complete spectral densities (continuous or sampled), we may consider a broader view of noise suppression, namely a decrease in the integrated spectral power over a bandwidth of frequencies. When the input signal is held constant ($\sigma^2 = 0$), the power spectrum for the continuous observation process Y_t will be flat with density $G(x_0)/2\pi$. When the input signal is allowed to fluctuate ($\sigma^2 > 0$), the spectrum adjusts to include the spectral density $h_H(\omega)$ in addition to the already existing white component corresponding to intrinsic noise. Because the intensity of the intrinsic noise is now averaged over the stationary distribution of X_t, the white component of the spectrum shifts to the level $\langle G(X_t)\rangle/2\pi$, which becomes a baseline for the spectral power at all frequencies. Let $I = (\omega_a, \omega_b)$ represent the spectral interval under consideration. When Eq. (49) holds and the baseline has decreased, the integrated spectral power on the interval I will decrease if

$$\int_{\omega_a}^{\omega_b} h_H(\omega)\,d\omega < \frac{G(x_0) - \langle G(X_t)\rangle}{2\pi}(\omega_b - \omega_a). \tag{55}$$

Since $h_H(\omega) \downarrow 0$ as $|\omega| \to \infty$, we will always be able to find a frequency interval (at least at high frequencies) where the integrated spectral power has decreased whenever the white spectral component has decreased. If we integrate the spectral density $h_H(\omega)$ over the entire real line, we recover the variance of $H(X_t)$,

$$\int_{-\infty}^{\infty} h_H(\omega)\,d\omega = \gamma_H(0) = \langle\widetilde{H}^2(X_t)\rangle, \tag{56}$$

since this amounts to computing the inverse Fourier integral for $u = 0$. Thus, the integrated contribution from $h_H(\omega)$ over any interval will always be less than the variance $\langle\widetilde{H}^2(X_t)\rangle$. If we consider the symmetric frequency interval $(-\omega_a, \omega_a)$, the integrated spectral density on this interval must decrease so long as

$$\omega_a > \pi\frac{\langle\widetilde{H}^2(X_t)\rangle}{G(x_0) - \langle G(X_t)\rangle}. \tag{57}$$

This is true even if the originally considered asymptotic noise has increased, $G_0 > G(x_0)$, which simply indicates that the spectral density begins above the original white spectral density and then falls below that level at higher frequencies. Note that if $G_0 < G(x_0)$, then the entire spectral density is below the original spectral density and thus the integrated spectral power will show a decrease over every interval.

For discretely sampled observations, we can similarly consider the integrated spectral power. A significant difference between the continuous and the sampled cases, however, is that the power spectrum for the sampled observation is defined on a compact interval while the spectrum for the continuous output signal has infinite support. Consequently, the spectral density $h_{H,\Delta}(\omega)$ will not vanish on this support so that we may not be able to find a frequency interval where the integrated spectral power has decreased even if the white spectral component has decreased, particularly when ρ is large. On the other hand, although the total power for the continuous process will be infinite, the total power for sampled observations is finite and given by

$$\int_{-\pi/\Delta t}^{\pi/\Delta t} h_{Y,\Delta}(\omega)\omega = \langle \widetilde{H}^2(X_k)\rangle + \langle \beta^2(X_k)\rangle, \tag{58}$$

the sum of the variance of $H(X_k)$ and the average variance of the added white noise. That is, the total spectral power is precisely equal to the variance of the output sequence. Thus, if

$$\langle \widetilde{H}^2(X_k)\rangle + \langle \beta^2(X_k)\rangle < \beta^2(x_0), \tag{59}$$

we would observe a decrease in the variance of the signal, which might be visualized as a narrowing of a histogram of observed values.

We remark in passing that the asymptotic noise levels G_0 and β_0^2 also have a well-established non-spectral interpretation. If we were to estimate the mean output signal H_0 using our observations, then G_0 would correspond to an asymptotic variance of that estimate based on continuous observations while β_0^2 would correspond to an asymptotic variance of the estimate based on sampled observations [8], such as one might compute using Markov chain Monte Carlo algorithms [10]. In particular, if we define the continuous and discrete sample means, respectively, as

$$\overline{Y}_T = \frac{1}{T}\int_0^T Y_t\,dt, \qquad \overline{Y}_N = \frac{1}{N}\sum_{k=0}^{n-1} Y_k, \tag{60}$$

then the variance of these estimators will asymptotically decrease as

$$\mathrm{Var}[\overline{Y}_T] \sim \frac{G_0}{T}, \qquad \mathrm{Var}[\overline{Y}_N] \sim \frac{\beta_0^2}{N}, \tag{61}$$

as $T \to \infty$ and $N \to \infty$, respectively.

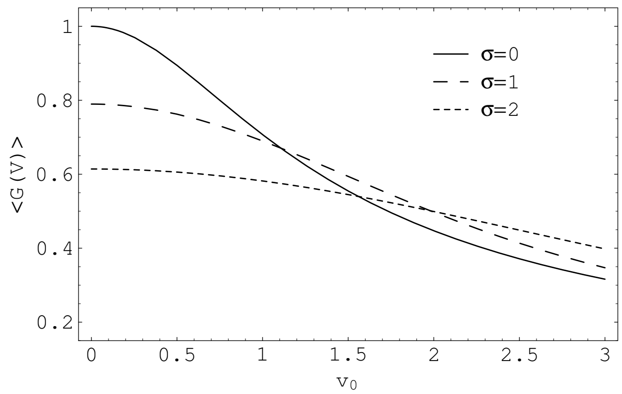

FIG. 1: The average intensity of intrinsic white noise, $\langle G(V_t)\rangle$, for various mean input signals (v_0) in the absence of state fluctuations ($\sigma = 0$) and for two levels of state fluctuations ($\sigma = 1$, $\sigma = 2$). Units are arbitrary.

V. EXAMPLE

We demonstrate these principles using an example proposed in [6] of a model for electrical conduction which displays saturation [11, 12]. The state $X = V$ corresponds to an input voltage. The observed current intensity has a mean characterized by the function

$$H(V) = \frac{V}{R(1+V^2)^{1/2}}, \tag{62}$$

and the intensity of the noise is characterized by

$$G(V) = \frac{Q}{(1+V^2)^{1/2}}, \tag{63}$$

where R and Q are constants. To conform to standard notation, we use V rather than X for this discussion. The parameters R and Q set the observation scale and the time scale, so that by rescaling the variables Y and t we may assume that $R = 1$ and $Q = 1$.

We first consider the effect of input fluctuations on the average intensity of the intrinsic output noise, $\langle G(V_t)\rangle$. Figure 1 plots the average intrinsic output noise intensity $\langle G(V_t)\rangle$ as a function of the mean input signal v_0 for two non-zero levels of input fluctuations ($\sigma = 1$ and $\sigma = 2$) as well as the original intensity of intrinsic noise ($\sigma = 0$), which simply corresponds to plotting $G(v_0)$. Averaging the intensity of intrinsic noise flattens and broadens the intensity profile as σ increases, so that the intensity profile completely vanishes in the limit as $\sigma \to \infty$. Whenever the intensity profile lies below the original intensity profile ($\sigma = 0$), the baseline level of the output signal power spectrum is decreased so that output noise is suppressed

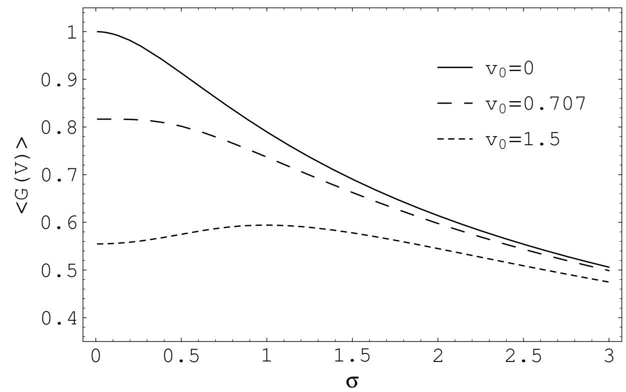

FIG. 2: The average intensity of intrinsic white noise, $\langle G(V_t)\rangle$, as a function of the size of input fluctuations (σ) for three values of the mean input signal ($v_0 = 0$, $v_0 = 1/\sqrt{2}$ and $v_0 = 1.5$). Units are arbitrary.

beyond some spectral frequency. Figure 2 shows the average intrinsic output noise intensity $\langle G(V_t)\rangle$ for fixed mean input signals v_0 as the size of the input signal fluctuations σ increases. For small values of σ, the behavior of $\langle G(V_t)\rangle$ is completely determined by $G''(v_0)$ using the perturbative approximation given by [6] that

$$\langle G(V_t)\rangle \approx G(v_0) + \frac{1}{2}\sigma^2 G''(v_0), \tag{64}$$

so that concavity initially determines whether the average noise intensity increases or decreases. When $v_0 < 1/\sqrt{2}$, the function G is concave down so that the average intensity initially decreases. Due to the specific example under consideration, this behavior continues for larger values of σ, so that the baseline of the power spectrum continually lowers. When $v_0 > 1/\sqrt{2}$, G is concave up so that the average intensity initially increases as the stationary distribution samples states (near $V = 0$) where the noise intensity is high. However, as σ continues to increase, the stationary distribution samples more extreme states where the intensity becomes asymptotically small. Consequently, the average intensity eventually decreases. Thus, although concavity determines whether the white component of the spectrum increases or decreases for small values of σ, it becomes irrelevant for predicting noise suppression for larger values of σ. The mean input level $v_0 = 1/\sqrt{2}$ corresponds to the point where $G''(v_0) = 0$, so that the fourth-order derivative, which is negative, determines that the average noise initially decreases proportionally to σ^4.

We next consider the power spectrum under conditions where noise suppression is possible, namely $\langle G(V_t)\rangle < G(v_0)$. Figure 3 plots typical power spectral densities corresponding

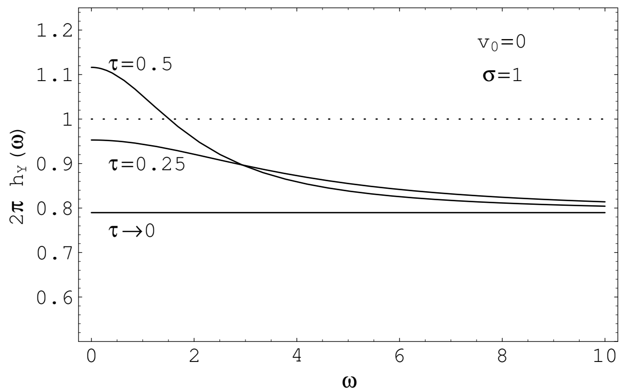

FIG. 3: The power spectral intensity $2\pi h_Y(\omega)$ as a function of spectral frequency (ω) for a mean input signal $v_0 = 0$ and input fluctuations of size $\sigma = 1$ where the input signal has autocorrelation time $\tau = 0.5$, $\tau = 0.25$, and in the limit $\tau \downarrow 0$. The reference white spectral density in the absence of input fluctuations is shown with a dotted line $(G(0) = 1)$. Units are arbitrary.

to $v_0 = 0$ and $\sigma = 1$. When the autocorrelation time for the input fluctuations is too large (e.g. $\tau = 0.5$), part of the spectral density is larger than the initial flat intensity at $G(0) = 1$. However, for sufficiently large frequencies ω, the spectral intensity decreases and approaches the new baseline $\langle G(V_t)\rangle < G(v_0)$, so that over sufficiently large spectral bandwidths, the integrated spectral power will decrease for any arbitrary autocorrelation time τ. But when the autocorrelation time is sufficiently small, the spectral intensity lowers and widens according to the scaling in Eq. (40) until it is uniformly smaller than the initial intensity (e.g. $\tau = 0.25$). The peak spectral intensity at $\omega = 0$ corresponds to the integrated scale G_0.

For those conditions where the average intensity of the intrinsic output noise decreases, the inequality of Eq. (50) determines a maximum autocorrelation time so that $G_0 < G(v_0)$. Recall that the suppression of the asymptotic scale of noise G_0 is equivalent to the power spectrum of the output noise lying completely below the original white spectrum in the absence of input fluctuations. In Fig. 3, this maximum time would correspond to that time τ where the spectral density is tangent to the original spectral density at $\omega = 0$. Figure 4 provides a contour plot of this maximum autocorrelation time as a function of v_0 and σ. The heavy contour line for $\tau = 0$ separates the region where noise suppression is possible from the region where noise invariably increases. We see that in this example the maximum autocorrelation time increases as the input fluctuations increase. Two factors make this possible. First, as the fluctuations σ^2 increase, the average intensity of the intrinsic noise

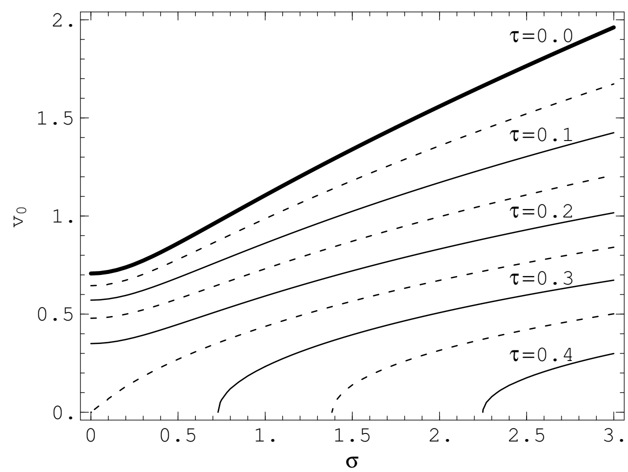

FIG. 4: A contour graph showing the largest possible autocorrelation time for the input fluctuations, τ, as a function of mean input signal v_0 and input fluctuation size σ such that noise is suppressed, $G_0 < G(v_0)$, as calculated in Eq. (50). Solid contours represent increments of $\Delta\tau = 0.1$, dashed contours represent increments of $\Delta\tau = 0.05$, and the coutour for $\tau = 0$ is shown in bold. Units are arbitrary.

$\langle G(V_t)\rangle$ decreases to zero. Second, the variation in the mean output signal $H(V_t)$ increases as σ^2 increases. However, since $H(V_t)$ is bounded between $+1$ and -1, added fluctuations at the input lead to smaller and smaller increases in the variation of $H(V_t)$. Consequently, for a given mean input v_0, the maximum autocorrelation time will be bounded. Considering the other direction, as $\sigma \downarrow 0$, the maximum autocorrelation time is determined from the previous perturbation results [6] to be

$$\tau = -\frac{G''(v_0)}{4H'(v_0)^2}, \tag{65}$$

so long as the condition $G''(v_0) < 0$ holds.

VI. CONCLUSION

In this article, we have considered noise suppression in the context of the power spectrum of an intrinsically noisy output signal for a system in which the correlations at the input decay exponentially. Because the intrinsic noise at the output depends on the input signal only in terms of the intensity, the autocovariance for the output signal and therefore the power spectrum both undergo a simple decomposition. This decomposition shows that the spectrum has a contribution arising from the average intensity of the intrinsic noise plus a component corresponding to the variation arising from the fluctuating signal itself. Consequently, noise suppression is possible only when the average intensity of the intrinsic noise decreases. That is, the increased fluctuations in the input signal must cause the system to visit more frequently those states where the intrinsic noise has low intensity such that the ergodic average of the intensity decreases.

Mathematically, the ergodicity of the input signal, rather than some nonlinearity in the system, plays the fundamental role in noise suppression. That is, according to the theory, there must be some input states where the intrinsic output noise is decreased. As the input signal samples its available phase space, it is driven into these low-noise states sufficiently often that the average intensity of noise decreases relative to an input signal constrained to a smaller phase space. In fact, the only strictly linear noise intensity profile $G(X)$ would be a constant, $G(X) = G$, since any nonzero slope would lead to negative intensities for some attainable state X. Nor, as we have shown, does concavity determine the existence of noise suppression except for sufficiently small fluctuations.

[1] R. D. Astumian, Science **276**, 917 (1997).

[2] Y. Okada and N. Hirokawa, Science **283**, 1152 (1999).

[3] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani, SIAM J. Appl. Math. **43**, 565 (1983).

[4] P. Hänggi, ChemPhysChem **3**, 285 (2002).

[5] D. F. Russell, L. A. Wilkens, and F. Moss, Nature **402**, 291 (1999).

[6] J. M. G. Vilar and J. M. Rubí, Phys. Rev. Lett. **86**, 950 (2001).

[7] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981).

[8] M. B. Priestley, *Spectral Analysis and Time Series* (Academic Press, London, 1981).

[9] E. Merzbacher, *Quantum Mechanics* (John Wiley & Sons, New York, 1970).

[10] A. Sokal, in *Functional integration (Cargèse, 1996)* (Plenum, New York, 1997), vol. 361 of *NATO Adv. Sci. Inst. Ser. B Phys.*, pp. 131–192.

[11] L. Reggiani, ed., *Hot Electron Transport in Semiconductors* (Springer-Verlag, Berlin, 1985).

[12] A. van der Ziel, *Noise in Solid State Devices and Circuits* (Wiley, New York, 1986).